Specifications for GIS Products and Deliverables
Last revision: August 18, 2005
Effective date: August 22, 2005

Purpose

These guidelines were developed for Geographic Information Systems (GIS) products created by, or submitted to, the North Coast and Cascades Network (NCCN) of the National Park Service (NPS). The primary goal is to improve GIS data quality and usability.

This document provides detailed guidance and specifications about the required software and formatting standards for I&M spatial data. These guidelines:
- provide direction to GIS users and data contributors about required GIS formats and products,
- promote consistency among GIS data formats,
- promote GIS product quality,
- state metadata requirements,
- facilitate GIS product identification,
- facilitate appropriate use of GIS products,
- facilitate links between spatial and non-spatial data, and
- minimize data errors and data format problems when GIS data are exchanged and used.

Scope and Applicability

These guidelines apply to all NCCN staff, contractors and cooperators responsible for generating and submitting data for permanent retention. Compliance with these guidelines is required for all I&M project databases (including geospatial databases), and for other projects as stipulated by project study plans, cooperative agreements, contracts or research permits. Specifically, all data that will be maintained in one of the NCCN Digital Libraries, or submitted to any of the NPS national databases/clearinghouses will be subject to these guidelines. Beyond these requirements, it is hoped that these guidelines will be adopted by others as a way to increase operational efficiency and compatibility among data sets.

Definitions and Acronyms

Coverage	The ESRI® ArcInfo® spatial vector data model. It generally represents a single set of topologically related (having spatial relationships between connecting or adjacent features) geographic objects such as roads, parcels, soil units or forest stands in a given area. These features are symbolized by points, lines, polygons, routes or regions. A coverage supports the georelational model - it contains both the spatial (location) and attribute (descriptive) data for geographic features. A coverage is comprised of a folder containing spatial data files and an "info" folder that contains tabular attribute data for each coverage (such as .pat, .aat files).
Database	A collection of data organized according to a conceptual structure describing data characteristics and relationships among corresponding entities. For example, a GIS database includes data about the position and characteristics of spatial features.

Data dictionary	A list of tabular fields and their properties. Each field is described and defined in terms of its data type, field size, format, domain values, validation rules, etc.
*Domain value*s	The range of values allowed in a field. For example, aspect can be any whole number between 1 and 360.
ESRI	Environmental Systems Research Institute. Company that creates and maintains GIS software such as ArcInfo and ArcGIS®.
FGDC	Federal Geographic Data Committee. Federal interagency organization that is developing the National Spatial Data Infrastructure (NSDI). The NSDI creates policies, standards, and procedures for geographic data production, metadata, and distribution.
Geodatabase	The ESRI name for 'geographic database'. The Geodatabase model is an ArcGIS version 8.0 and above data format. A geodatabase represents geographic features and attributes as objects and is hosted inside a relational database management system (DBMS).
GIS	Geographic Information Systems. A computer system for creating, storing, checking, integrating, manipulating, analyzing, and displaying spatial data. GIS software typically used by NPS includes ArcInfo, ArcView®, ArcGIS (ArcMap, ArcCatalog, ArcToolbox), and ArcIMS®.
Metadata	Documentation describing a dataset. Typically describes purpose, geographic extent, content, condition, distribution, and contact information.
NBII	National Biological Information Infrastructure. Collaborative program that developed the "Biological Data Profile" portion of the FGDC metadata standards. NBII extends FGDC standards to include information about biological aspects of data, such as taxonomy.
NCCN	North Coast and Cascades Network http://www1.nature.nps.gov/im/units/nccn
NPS	National Park Service
Primary key	A tabular field that uniquely identifies each record in a relational table. It can either be a normal attribute that is guaranteed to be unique or it can be generated by the DBMS (such as a globally unique identifier, or GUID, in Microsoft SQL Server). Primary keys may consist of a single attribute or multiple attributes in combination.
Raster	A data model that stores data by rows and columns, forming cells. Each cell is a discrete, uniform, geographically located unit containing a single data value that represents a feature or characteristic. ESRI raster datasets are called grids.
Shapefile	An ESRI GIS data format that stores non-topological geometry (see definition of coverage) and attribute information for the spatial features. The geometry for a feature is stored as a shape comprised of a set of vector coordinates. Shapefiles can support point, line and area features.
Vector	A data model that stores points, lines, or polygons. The spatial features' locations are defined by x, y coordinates (points) or a series of x, y coordinates connected by linear segments (lines, polygons).
XML	Extensible mark-up language.

Overview

These guidelines outline required documentation and formats for spatial data submitted to the NCCN. Spatial data product specifications, NCCN coordinate system and datum standards, quality assurance and quality control considerations, and metadata requirements are discussed. A list of products required for submission to NCCN is provided in Table 1.

Table 1. Required spatial products and associated information for delivery to NCCN:

Required
- Spatial data

- Associated attribute information (Access Database, dBase files, INFO files, etc.)

- FGDC-compliant metadata (XML format) or descriptive documentation text file

If Applicable
- Associated map legends, specialized map symbology

- Linked documents, graphics, digital photographs

Procedures and General Requirements

NCCN GIS Software

Products submitted to NCCN should conform to the latest version of software used by the NPS. NCCN primarily uses ESRI GIS products on Microsoft® Windows operating systems. ESRI products generally used by NPS include ArcInfo, ArcView, and ArcGIS (ArcMap, ArcCatalog, and ArcToolbox).

Spatial Data Products Specifications

GIS Formats

Acceptable spatial data formats are ArcInfo coverages or grids in export file format (.e00) and shapefiles (.shp, .shx, .dbf, and .prj at a minimum). NCCN has not yet transitioned to the ESRI geodatabase format due to resource constraints, including training and development time, and the cost of licensing for enterprise-level database administration. This option also does not fit as easily into current work flow practices by network staff. Geodatabase format will be accepted, but is not preferred.

Non-GIS Formats

If no GIS support is readily available, then spatial data may be provided in tabular or map formats that can be converted into GIS products by NCCN GIS specialists.

Maps

Hardcopy or digital maps may be submitted. Typically, these consist of a background map with graphics drawn on top to represent the spatial features of interest. Stable base transparencies are recommended for projects involving a large number of features. Using GIS software, these graphics are then transcribed (digitized) onto a computer and are tied to geographic locations thus becoming spatial data with real-world coordinates. The features can then be linked to tabular data containing attribute information via a common field (primary key).

A United States Geological Survey (USGS) 7.5 minute (1:24,000 scale) quadrangle map is a good example of a standard product that can be used as a background for graphics (points, lines, or polygons) representing spatial data. These maps provide sufficient detail from which data can be accurately transcribed. Contact park GIS specialists for information about other background map options, including digital raster graphics (DRGs), digital orthophoto quadrangles (DOQs) and quarter-quadrangles (DOQQ's), air photos, etc.

Minimum map requirements:
- Maps must be clear and legible.
- Maps should include a title, a graphic scale, a north arrow, and a legend.
- Maps must be at an appropriate scale for the context of the spatial data being displayed.
- Maps must include labeled graticules or at least four corner tic marks showing coordinates preferably in Universal Transverse Mercator (UTM), or other projected coordinate systems.
- Spatial features should be labeled in a manner that matches a report or submitted tabular data.
- Digital maps may be submitted in any standard image format such as JPEG (.jpg), Microsoft Windows bitmap (.bmp), or Tagged Image Format (.tiff), or they may be inserted into a Word (.doc), PowerPoint (.ppt), or Adobe Acrobat (.pdf) document.
- Depending on project objectives, spatial data on maps should be accompanied by tabular data. See the specifications below for tabular data formats.
- Maps must be accompanied by a text file containing information on map author, date of map creation and source of data. Digital maps should also include a source pathway.

Tabular Spatial Data

GIS layers can be created from tabular data. Point feature data will be accepted in ArcInfo generate, Excel spreadsheet, dBase, MS Access table, or comma delimited text file formats. See Appendix 1 for ArcInfo generate file format example for points. Spreadsheet data to create a point feature GIS layer must, at a minimum, have:
- one column containing a unique record number (GIS_Loc_ID)
- one column containing X coordinates (defined as numeric)
- one column containing Y coordinates (defined as numeric)
- one column defining estimated horizontal accuracy in a numeric format (specify measure units in column heading, e.g. 'elev_m' for elevation in meters)

Line, polygon, and grid data will be accepted in ArcInfo generate format. See Appendix 1 for generate file format examples. See Database Template (NCCN 2005b) and GIS Naming Conventions (NCCN 2005d) for more detailed information about standard fields and formats for NCCN GIS data.

Coordinate System

The current NCCN coordinate system standard is Universal Transverse Mercator (UTM), Zone 10 Measure units for the UTM projection system are meters.

Datum

The current NPS standard is North American Datum (NAD) 1983.

Ellipsoid

The current NPS standard is the Geodetic Reference System (GRS) 80.

Spatial Scale

A spatial scale appropriate to accomplishing project spatial objectives should be used. Consult with NCCN GIS staff to discuss spatial scale details.

Spatial Accuracy

Spatial accuracy must be appropriate to project objectives. Consult with NCCN GIS staff and project managers to discuss accuracy requirements and appropriate protocols to meet those requirements.

Tabular Attribute Data

Spatial features rarely consist of simply coordinate information. Most spatial features have associated data, called attributes (also called "fields" in a table), that describe characteristics of those features. A GIS allows users to query, analyze, and display spatial data based on these attributes. Spatial features can be associated with tabular data a number of ways. Spatial features can link to external databases using Open Database Connections (ODBC), using join and relate GIS functions, or using specialized software such as ArcGIS extension programs. Refer to Spatial Data Integration (NCCN 2005h) for more information.

Data Format

Tabular attribute data should be submitted in comma or space delimited text files, Microsoft Excel spreadsheet files, Microsoft Access tables, or dBase files.

Attribute Requirements

Tabular attribute data should:
- be related to spatial coordinate information in either the same or a different delimited text, Excel, Access, or dBase file. For point features, these data can be either presented in columns alongside coordinate data or provided in a separate table or file with a primary key specified. For all other feature types, attribute data should be provided in a separate table.
- have an explicitly defined primary key (GIS_Loc_ID) that relates spatial coordinates to attribute data in a separate file. See Database Template (NCCN 2005b).
- be in the most normalized form possible. See Database Template (NCCN 2005b).
- conform to GIS field naming standards such as having ten or fewer characters in each field name for attributes that are to be spatially queried or analyzed in GIS. See GIS Naming Conventions (NCCN 2005d).
- be accompanied by a data dictionary document (this is part of FGDC metadata) that states attribute definitions, such as field format and size and domain values. For example, a field containing aspect would be defined as numeric with a domain of values between 1 and 360; this avoids confusion about whether 0 or 360 designates north. See Database Documentation (NCCN 2005a, in development).

Attribute Accuracy

There should be a minimum attribute accuracy of 95%. Consult GIS Development Guidelines (NCCN 2005c, in development) for appropriate protocols to meet data quality objectives.

Associated Symbology, Documents, Graphics, Digital Photographs

Associated files should be stored with the data set. Unique map symbol sets and legend files (ex. _.avl) should be included with GIS data. Graphics, digital photos, or documents linked to GIS layers should be organized into a folder along with a text document describing file naming conventions and relationships to spatial data (i.e. dataset, table, and field names). Links to these associated files should be stored in GIS attribute tables using relative pathnames. Consult ArcView and ArcMap help files for information on linking text and image files with spatial data. See GIS Naming Conventions (NCCN 2005d) and Database Documentation (NCCN 2005a, in development) for further guidance.

Quality Assurance/Quality Control

A small subsample, approximately 10% of the data, should be entered into the GIS and tested before all data are entered. Coordinate and attribute data should be verified and validated periodically as they are entered. Spatial accuracy should be evaluated at a scale appropriate at which data were collected and intended to be used. Data *verification* is checking that digitized data match source data. Data *validation* is checking that data are free from logical errors. Documentation of the quality assurance and control standards used in producing the information must be formally documented in metadata or in a descriptive document file. See Quality Assurance Certification (NCCN 2005f, in development) and GIS Development Guidelines (NCCN 2005c, in development) for details.

Metadata

Whenever possible, spatial data should be submitted in a GIS format with accompanying FGDC compliant metadata, preferably in XML format (.xml). FGDC metadata, in conjunction with NBII metadata where appropriate, will provide the most complete documentation for a spatial dataset. For more information on FGDC metadata see http://www.fgdc.gov/metadata/metadata and on NBII metadata see http://www.nbii.gov/datainfo/metadata. See Metadata Guidelines (NCCN 2005e, in development) for help on metadata creation for NCCN products.

In the absence of FGDC metadata, there must be a descriptive document (Word, PDF, or Readme.txt file). This descriptive document must specify at a minimum:
- List of CD, DVD, or .zip file contents
- GIS data layer creation date and version
- Date of data collection
- Contact information (name, phone number, address, agency or organization) for person responsible for the data
- GIS data format(s)
- Short description of each GIS layer and associated files (tables, images, databases, etc.)
- Coordinate system and measurement units
- Datum
- Ellipsoid
- Data set purpose
- Spatial scale or resolution (in the case of raster data)
- Horizontal accuracy estimates

- Attribute accuracy level (percent)
- Source(s) from which dataset was derived
- Methods describing how dataset was created
- Indication if the data contain sensitive information
- Data dictionary (see Database Documentation (NCCN 2005a, in development))
- Corresponding thematic information such as legend files, specialized symbology, hyperlinked images, or documents.
- Corresponding relational data (such as MS Access databases, INFO files, or dBase files)
- Primary key identification between GIS layers and corresponding relational data

Spatial Data Products Delivery

Spatial products are to be delivered to one of the NCCN's GIS specialists or data managers.

Data should be digital whenever possible. Digital data can be compressed using WinZip. Files should be delivered on CD or DVD. Compact discs should be in CD-R and ISO 9660 format to allow cross-platform use. Alternatively, File Transfer Protocol (ftp) sites or e-mail may be used to deliver products. FTP sites and e-mail should not be used for data containing sensitive information (see Sensitive Information Guidelines (NCCN 2005g). E-mail should not be used for files larger than 5 MB.

Refer to Table 1 for summary of required spatial products delivered to NCCN.

Responsibilities

- Project managers, data managers, and NPS cooperators and contractors who will create spatial data or use GIS capabilities are responsible for consulting with NCCN GIS staff during data design and development.
- NCCN GIS staff members are responsible for providing support to all parties requesting information about I&M GIS data development and delivery within the NCCN.
- NCCN GIS Specialists will verify that submitted GIS layers adhere to formats outlined in this document.

Credits

Parts of these guidelines contain material from:
- NPS GIS Program. 2005. NPS GIS Data Specifications for Resource Mapping, Inventories and Studies. USDI National Park Service. Available at: http://science.nature.nps.gov/nrgis/standards/docs/GISSpecs31105_final.pdf
- Colonial NHP. 1998. SOP for Themes, Attribute Tables, and Map Design. USDI National Park Service.

Recommended Citation

North Coast and Cascades Network – National Park Service. 2005. GIS Product Specifications. USDI National Park Service.

Contact Information

Mount Rainier National Park   Darin Swinney, GIS Coordinator
Tahoma Woods – Star Route   360-569-2211 ext. 3378
Ashford, WA 98304   darin_swinney@nps.gov

North Cascades National Park Service Complex   Anne Braaten, GIS Coordinator
810 State Route 20   360-856-5700 ext. 238
Sedro-Woolley, WA 98284-1239   anne_braaten@nps.gov

Olympic National Park   Roger Hoffman, GIS Coordinator
600 East Park Av.   360-565-3062
Port Angeles, WA 98362-6757   roger_hoffman@nps.gov

Ebey's Landing National Historical Reserve   Please refer to contacts for North Cascades NP

Fort Vancouver National Historic Site   Please refer to contacts for Mount Rainier NP

Lewis and Clark National Historical Park   Please refer to contacts for Mount Rainier NP

San Juan Island National Historical Park   Please refer to contacts for North Cascades NP

Reference Documents

Related Guidance
- North Coast and Cascades Network – National Park Service. 2005a. Database Documentation. USDI National Park Service. Available at: http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm (in development).
- North Coast and Cascades Network – National Park Service. 2005b. Database Template. USDI National Park Service. Available at: http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm.
- North Coast and Cascades Network – National Park Service. 2005c. GIS Development Guidelines. USDI National Park Service. Available at: http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm (in development).
- North Coast and Cascades Network – National Park Service. 2005d. GIS Naming Conventions. USDI National Park Service. Available at: http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm.
- North Coast and Cascades Network – National Park Service. 2005e. Metadata Guidelines. USDI National Park Service. Available at: http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm (in development).
- North Coast and Cascades Network – National Park Service. 2005f. Quality Assurance Certification. USDI National Park Service. Available at: http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm (in development).
- North Coast and Cascades Network – National Park Service. 2005g. Sensitive Information Guidelines. USDI National Park Service. Available at: http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm.
- North Coast and Cascades Network – National Park Service. 2005h. Spatial Data Integration. USDI National Park Service. Available at: http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm.

Tutorials
- Federal Geographic Data Committee (FDGC) tutorial information http://www.fgdc.gov/metadata/metatut.html
- National Biological Information Infrastructure (NBII) training information http://www.nbii.gov/datainfo/metadata/training/ttt/index.html

Other Citations and References
- Boetsch, J.R., B. Christoe, and R.E. Holmes. 2004. Draft data management plan for the North Coast and Cascades Network Inventory and Monitoring Program. USDI National Park Service. Port Angeles, WA. Available at: http://www1.nature.nps.gov/im/units/nccn/dm_docs.htm
- NPS GIS Program. 2005. NPS GIS Data Specifications for Resource Mapping, Inventories and Studies. USDI National Park Service. Available at: http://science.nature.nps.gov/nrgis/standards/docs/GISSpecs31105_final.pdf

Revision History

Revision Date	Description of Change	Author	Effective Date
Mmm dd, yyyy		Full name	Mmm dd, yyyy

Appendix 1. Generate File Formats for Point, Line, Polygon, and Grid Data

<u>Point Feature Data</u>

An ArcInfo point generate file consists of a comma delimited text file where each entry represents a coordinate pair with an identification number (ID). The file's final line contains "end" to mark the end of the generate file.

For example:
ID,x,y
ID,x,y
ID,x,y
end

<u>Line Feature Data</u>

For line feature data, coordinates must be in a comma delimited text file that conforms to the ArcInfo line generate format. This format provides a line identification number (ID) followed by a series of coordinate pairs (x, y) that spatially define each vertex of the line. Each line's end-point is indicated with "end" on the line following the final vertex coordinate pair. The file's final line contains "end" to mark the end of the generate file.

For example:
ID,x,y
x,y
x,y
end
ID,x,y
x,y
end
end

<u>Polygon Feature Data</u>

For polygon feature data, vertex coordinates must be in a comma delimited text file that conforms to the ArcInfo polygon generate format. This format provides a label identification number (ID) for each polygon followed by coordinate pairs that represent polygon vertices. The final coordinate pair needs to be the same as the first coordinate pair to mark the closing of the polygon. The series of coordinate pairs representing a polygon is finished with "end" after the line containing the final vertex coordinate pair. The file's final line contains "end" to mark the end of the generate file.

For example:
ID,x,y
X,y
X,y
X,y
end
ID,x,y
X,y
X,y
X,y

X,y
end
end

Grid Feature Data

Grids can be generated using a text (ASCII) file. The ASCII file contains header information followed by
cell values. The header information states the number of columns, the number of rows, the lower left cell
corner or lower left cell center coordinates, and the cellsize. Cell values should be delimited by spaces.
Carriage returns are not necessary at the end of each row in the grid because the number of columns in the
header is used to determine when a new row begins. The number of cell values must be equal to the
number of rows times the number of columns, or an error will be returned (ESRI ArcInfo Help).

The file format is:
<ncols xxx>
<nrows xxx>
<xllcenter xxx | xllcorner xxx>
<yllcenter xxx | yllcorner xxx>
<cellsize xxx>
{nodata_value xxx}
row 1
row 2
.
.
row n

Within this format, xxx is a number. Row 1 of the data is at the top of the grid, row 2 is just under row 1
and so on. There is an optional keyword, "nodata_value" that defaults to -9999. The nodata_value is the
value in the ASCII file to be assigned to those cells whose true value is unknown. In the grid they will be
displayed as "NODATA" (ESRI ArcInfo Help).

For example:
ncols 480
nrows 450
xllcorner 378923
yllcorner 4072345
cellsize 30
nodata_value -32768
43 2 45 7 3 56 2 5 23 65 34 6 32 54 57 34 2 2 54 6
35 45 65 34 2 6 78 4 2 6 89 3 2 7 45 23 5 8 4 1 62...